Identification of genetic factors contributing to disease susceptibility in Multiple Sclerosis exemplified by a candidate gene search in the *CD28/CTLA4/ICOS* gene region

By
Åslaug Rudjord Lorentzen

for
The Medical Student Research Program
Faculty of Medicine
University of Oslo

Tutor Hanne Flinstad Harbo
Co-tutor Anne Spurkland

Institute of Immunology
Faculty Division Rikshospitalet
University of Oslo

2005

CONTENTS

1 ACKNOWLEDGEMENTS

The work presented in this study has been carried out at the Institute of Immunology (IMMI), Faculty Division Rikshospitalet, University of Oslo (UiO), Norway supported by The Medical Student Research Program, Faculty of Medicine, UiO (project number;131409/000075/410993) during the years 2003-2005. The project has also received funding from the Norwegian Research Council (project number 154888/V40), Odd Fellow MS society, The Multiple Sclerosis Society of Norway and stud.med. Morten Dedekam Harboe's Fund.

First of all I would like to thank my supervisor Hanne Flinstad Harbo. Without your enthusiasm and "tough love", this work had never been possible. You have been a great inspiration for me and all your knowledge have you kindly shared.
I am grateful to Professor Frode Vartdal, who was the first person which introduced me to the immunological research. A special thank to my co-supervisor Anne Spurkland and Benedicte A. Lie for great guidance and support.
Thank to Erik Thorsby, head of the Institute of Immunology and the Immunogenetic group, and all my co-workers at IMMI for making a good scientific and social environment. The excellent technical assistance of Karen-Marie Heinz, Kristina Narvhus, Ingebjørg Knutsen Heitman, Monica Hals and Wenche Scheel Hamang is greatly acknowledged. The Norwegian Bone Marrow Registry is thanked for collaboration in establishment of the control material and collaboration on HLA typing.

Outside the Institute I want to thank Per Ekstrøm at Department of Surgery, Institute for Cancer Research, The Norwegian Radium Hospital, Oslo, Norway and Stephen Sawcer at Department of Neurology, University of Cambridge, Cambridge, UK for their supportive collaboration. Elisabeth G Celius at Department of Neurology at Ullevål University Hospital, is warmly thanked, especially for invaluable collaboration concerning the collection of the MS material and collection and analysis of the clinical material. Kjell-Morten Myhr at Department of Neurology at Haukeland University Hospital is also thanked for contributing with additional MS samples.

Finally, I want to thank The Medical Student Research Program head and course organizer associate Professor Jarle Breivik and the course administrator senior executive officer Maje Siebke. They have always been there for questions and good support.

The thesis is dedicated to my husband and best friend Øyvind.

2 ABBREVIATIONS

APC	antigen presenting cell
BBB	blood brain barrier
BCR	B cell receptor
bp	base pair
CD	cluster of differentiation molecule
cM	centiMorgan
CNS	central nervous system
CTLA-4	cytotoxic T lymphocyte associated antigen 4
DCE	denaturant capillary electrophoresis
DNA	deoxyribonucleic acid
DZ	dizygotic twins
EDSS	expanded disability status scale
FS	functional systems
FSS	functional system score
HLA	human leukocyte antigen
ICAM	intracellular adhesion molecule
ICOS	inducible co-stimulatory molecule
Ig	immunoglobulin
IL	interleukin
JRA	juvenile rheumatoid arthritis
k	kilo
LD	linkage disequilibrium
MHC	major histocompatibility complex
MR	magnetic resonance imaging
MS	Multiple Sclerosis
MZ	monozygotic twins
OR	odds ratio
PCR	polymerase chain reaction
PP-MS	primary progressive Multiple Sclerosis

RFLP	restriction fragment length polymorphism
RR-MS	relapsing remitting Multiple Sclerosis
SH2D2A	SH2 domain protein 2A gene
SNP	single nucleotide polymorphism
TCR	T cell receptor
Th cell	T helper cell
TNF	tumor necrosis factor
TSAd	T cell specific adaptor protein
VNTR	variable number of tandem repeats

3 LIST OF PUBLICATIONS

The paper focused at in this thesis:

Lorentzen AR, Celius EG, Ekstrom PO, Wiencke K, Lie BA, Myhr KM, Ling V, Thorsby E, Vartdal F, Spurkland A, Harbo HF. *Lack of association with the CD28/CTLA4/ICOS gene region among Norwegian multiple sclerosis patients*. Accepted 2005 June 7 in J Neuroimmunol.

Other related papers:

Smerdel A, Dai KZ, **Lorentzen AR**, Flato B, Maslinski S, Thorsby E, Forre O, Spurkland A. *Genetic association between juvenile rheumatoid arthritis and polymorphism in the SH2D2A gene*. Genes Immun. 2004 Jun;5(4):310-2.

Harbo HF, Lie BA, Sawcer S, Celius EG, Dai KZ, Oturai A, Hillert J, **Lorentzen AR**, Laaksonen M, Myhr KM, Ryder LP, Fredrikson S, Nyland H, Sorensen PS, Sandberg-Wollheim M, Andersen O, Svejgaard A, Edland A, Mellgren SI, Compston A, Vartdal F, Spurkland A. *Genes in the HLA class I region may contribute to the HLA class II-associated genetic susceptibility to multiple sclerosis.* Tissue Antigens. 2004 Mar;63(3):237-47.

4 GENERAL INTRODUCTION

4.1 Multiple Sclerosis

Multiple Sclerosis (MS) is a chronic demyelinating disease of the central nervous system (CNS) affecting young adults. MS is one of the most frequent causes of major neurological disability among people in the western world. In a Norwegian study including patients from the Oslo area the prevalence among patients of native Norwegian ancestry was found to be 136/100 000[1]. This is in accordance to earlier reports that the south-eastern part of Norway is among the regions with the highest MS prevalence in the country and in the world[2,3]. In Nord-Trøndelag County a high prevalence is also shown; 163.6/100 000 which is among the highest ever reported in Norway[4] (Figure 1).

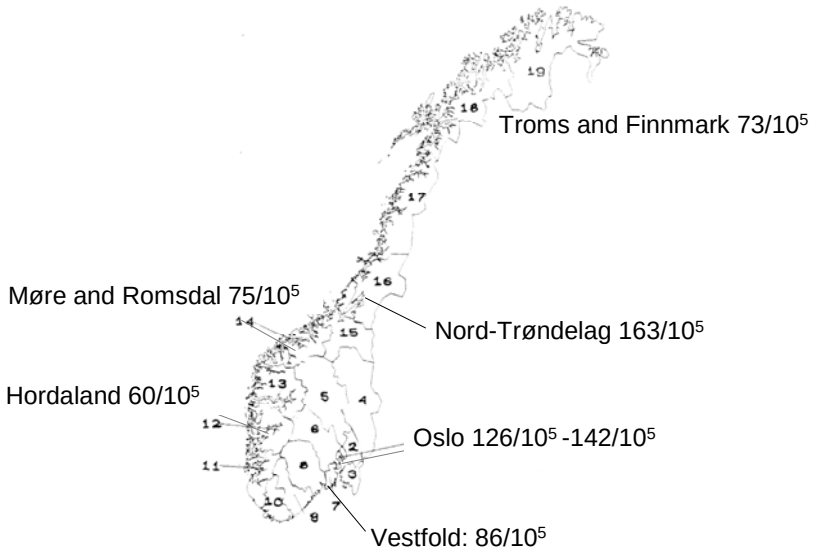

Figure 1: Prevalence of Multiple Sclerosis (MS) in Norway

The majority of the patients are affected by relapsing remitting MS (RR-MS), which is characterized by episodic attacks of neurological deficits, followed by periods of remission. The rate of disease progression may vary, but almost all RR-MS patients develop a secondary progression disease course, so-called secondary progressive (SP) MS[5]. The remaining 20 % of the MS patients develop a primary progressive disease course, so-called PP-MS. These do not experience acute relapses, but develop the

symptoms of neurological deficits progressively. Females are affected twice as often as males, and the mean age at onset is usually between the age of 20 and 40 years[6].

The symptoms of MS depend on the location of the inflammatory and demyelinating lesions within the CNS, and are therefore very variable as shown in Table 1.

LESION LOCATION:	SIGNS/SYMPTOMS:
Cerebrum & Cerebellum	Balance problems, speech problems, coordination, tremor
Motor nerve tracts	Muscle weakness, spasticity paralysis, vision problems, bladder and bowel problems
Sensory nerve tract	Altered sensation, numbness, prickling, burning sensation

Table 1: Possible location of the Multiple Sclerosis (MS) lesions and symptoms and signs of the disease. From http://www.msfacts.org/info/info_symptoms.html

In RR-MS the disease begins with an acute or subacute onset of neurological abnormalities, a so-called MS attack, MS relapse or MS schub. Early symptoms may include numbness, paresthesia, paresis, double vision, optic neuritis, and ataxia and bladder control problems. Also more unspecific symptoms as depression, emotional lability, fatigue and pain are commonly seen.

Already 100 years ago Charcot described the clinical and pathological characteristics of MS[5]. Still the MS diagnosis is mainly based on the patient history and clinical examination. In 1983 Poser et al published a set of criteria for research purposes[6]. These criteria are now used also for clinical practice. The MS-diagnosis is based on the history of at least two attacks affecting more than one anatomical site in CNS (so-called dissemination in time and place). In 2001 the MS criteria were revised by "The International Panel on MS Diagnosis" [7]. Magnetic resonance imaging (MR) was then integrated in the clinical diagnostic criteria.

Expanded Disability Status Scale (EDSS) is a classification system of MS[8]. The EDSS quantifies disability in eight Functional Systems (FS) and allows neurologists to assign a Functional System Score (FSS) in each of these. The FS are as follows; pyramidal, cerebellar, brainstem, sensory bowel and bladder, visual, cerebral and other. The EDSS score 0.0 means normal neurological examination, whereas score 10.0 is death due to MS. EDSS steps between 1.0 and 4.5 refer to people with MS who are fully ambulatory. EDSS steps between 5.0 and 9.5 are defined by the impairment to ambulation. The EDSS can be used by Neurologist to follow the progression of MS disability, evaluate treatment results and for scientific purposes. For more details see the Table 2 below.

0.0	Normal neurological examination
1.0	No disability, minimal signs in one FS
1.5	No disability, minimal signs in more than one FS
2.0	Minimal disability in one FS
2.5	Mild disability in one FS or minimal disability in two FS
3.0	Moderate disability in one FS, or mild disability in three or four FS. Fully ambulatory
3.5	Fully ambulatory but with moderate disability in one FS and more than minimal disability in several others
4.0	Fully ambulatory without aid, self-sufficient, up and about some 12 hours a day despite relatively severe disability; able to walk without aid or rest some 500 meters
4.5	Fully ambulatory without aid, up and about much of the day, able to work a full day, may otherwise have some limitation of full activity or require minimal assistance; characterized by relatively severe disability; able to walk without aid or rest some 300 meters.
5.0	Ambulatory without aid or rest for about 200 meters; disability severe enough to impair full daily activities (work a full day without special provisions)
5.5	Ambulatory without aid or rest for about 100 meters; disability severe enough to preclude full daily activities
6.0	Intermittent or unilateral constant assistance (cane, crutch, brace) required to walk about 100 meters with or without resting
6.5	Constant bilateral assistance (canes, crutches, braces) required to walk about 20 meters without resting
7.0	Unable to walk beyond approximately five meters even with aid, essentially restricted to wheelchair; wheels self in standard wheelchair and transfers alone; up and about in wheelchair some 12 hours a day
7.5	Unable to take more than a few steps; restricted to wheelchair; may need aid in transfer; wheels self but cannot carry on in standard wheelchair a full day; May require motorized wheelchair
8.0	Essentially restricted to bed or chair or perambulated in wheelchair, but may be out of bed

	itself much of the day; retains many self-care functions; generally has effective use of arms
8.5	Essentially restricted to bed much of day; has some effective use of arms retains some self care functions
9.0	Confined to bed; can still communicate and eat.
9.5	Totally helpless bed patient; unable to communicate effectively or eat/swallow
10.0	Death due to MS

Table 2: Kurtzke Expanded Disability Status Scale (EDSS) for Multiple Sclerosis (MS) patients.

From http://www.mult-sclerosis.org/expandeddisabilitystatusscale.html

The prognosis of MS is highly variable and difficult to predict. A review from 2001[9] analyzed possibilities for prediction of a benign course of MS. Especially, onset with optic neuritis, onset before the age of 40 years, absence of pyramidal signs at presentation, duration of first remission more than one year, and only one exacerbation in the first 5 years after onset of MS, were associated with a benign course, which was seen in 26.7% of the cases.

4.2 The immune system

The function of the immune system is to recognize and eliminate invading foreign organisms. The immune system must therefore distinguish self from non-self and dangerous from non-dangerous. Immune responses are classified in several ways: as primary (induced by first contact with antigen) or secondary (renewed contact with antigen); as humoral immunity mediated by immunoglobulins (Igs) produced by B-lymphocytes (B-cells) or as cellular immunity mediated by cytotoxic T-lymphocytes (T-cells). ”Adaptive immunity” gives a specific lifelong immunity. “Innate immunity” is not specific and consist of mainly macrophages and neutrophiles which are the first line of defense against foreign microorganism [10,11].

4.2.1 Components of the immune system

Antigens are substances with a molecular surface structure that trigger an immune response; i.e. production of antibodies or involvement in the specific antigen-antibody reaction. Antigens are often classified according to source, as exogenous or endogenous.

Major histocompatibility complex (MHC) molecules are surface molecules encoded by the class I or class II regions of the MHC gene, in humans known as the human leukocyte antigen (HLA) molecule and gene. Class I molecules are found on all cells and are specialized to present antigen-peptide synthesized within the host cells. The molecule can therefore present viral and tumor antigens. MHC class II molecules are found mainly on antigen presenting cells (APCs). They are specialized to present antigen-peptide derived from outside the host cells, such as bacterial fragments and other extracellular antigens.

B-cells are lymphocytes that express the B-cell receptor (BCR). The BCR is a membrane-bounded immunoglobulin (IgM) an antibody, and serve as the cell’s receptor for antigen. On activation by an antigen and with stimulatory signals and growth factors released locally by T helper (Th) cells, the B-cell differentiates into a so-called plasma cell. The plasma cells produce antibody molecules of the same antigen specificity as the receptor on the B-cells. The secreted antibodies have the capacity to bind to the inducing antigens and label them for further phagocytosis by macrophages.

T-cells are lymphocytes that usually express the cluster of differentiation 3 (CD3) molecules. Almost every T-cell expresses T-cell receptors (TCRs). The TCR is an

antibody-like molecule located on the surface membrane on the T-cell. When the antigen-peptide is presented in the cleft of either MHC class I or class II molecules, the T-cells can recognize the antigen, and this stimulates to further signal in the T-cells. All activated T-cells secrete interleukin-2 (IL-2), which stimulate to production of IL-2 receptors and proliferation of T-cells.

The T-cells fall into two classes that perform different but overlapping functions. The two classes are distinguished by the expression of the surface markers CD4 and CD8. These adhesion molecules on the T-cell surface act as co-receptors in the interaction between the TCR and the antigen-peptide-MHC-complex on the APC. CD4 is expressed on approximately 60 % of mature T-cells, whereas CD8 is expressed on about 30% of the T-cells. The CD4/CD8 ratio in normal healthy individuals is therefore about 2:1.

The CD4+ T-cells are "helper" T-cells (Th-cells). Their main function is as regulatory cells which modulate the function of B-cells and other T-cells through the secretion of cytokines. Further, the CD4+ cells can be divided into two subgroups based on the cytokine profile and predominant function. Type 1 CD4+ cells help directly in the cell-mediated immune response, producing interferon-gamma (IFN-γ), tumor necrosis factor alpha and beta (TNF-α, TNF-β) and IL-2 which are pro-inflammatory cytokines. The type 2 CD4+ cells secreted IL-4, IL-5, IL-6, and IL-10 and have an opposite effect to Th1. They participate in the immediately hypersensitivity reactions and are important in the B-cell antibody-mediated immunity[10]. The CD4+ T-cells recognize antigen presented in the MHC class II molecules, and the MHC class II molecules are found mainly on APCs.

The CD8+ T-cells or cytotoxic T-cells recognize antigen presented on the class I MHC molecules expressed at all cells. They act as killer's, and kill "infected host cells", which express the antigen-peptide in the MHC class I molecule on the cell surface.

4.2.2 The adaptive immune system

The adaptive immune system is specific and has "memory". The memory is expressed as a rapid and strong secondary response to an antigen that earlier has given a primary immune response. In the primary immune response memory cells have developed from the clonally expanded B- or T-cells. The memory cells are long lived and they are in

continuously recirculation in the blood waiting for new attacks from earlier detected antigens.

A complete activation of the T-cell requires two signals of which the first is antigen-specific and delivered through the TCR. This first signal is required for the naïve T-cells to proliferate and differentiate, and is the interaction between the TCR on the T-cell and the antigen-peptide-MHC-complex on the APC (Figure 2).

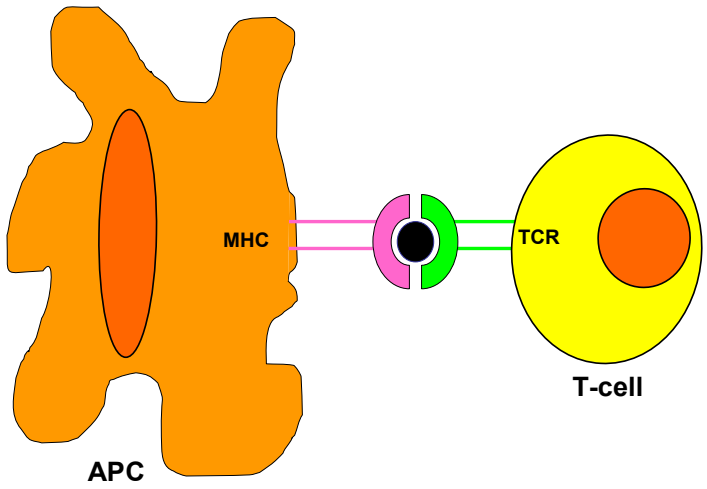

Figure 2: The primary interaction between the MHC molecule on the antigen presenting cell (APC) and the T-cell receptor (TCR) on the T-cell. The antigen is coloured black on the figure.

The second or costimulatory signal (Figure3) is not antigen-specific and is transmitted through interaction between other surface molecules on the T-cell and the APC. This signal is essential for the functional response in the T-cells, and is transduced through interaction between the cluster of differentiation 28 (CD28) molecules or cytotoxic T lymphocyte antigen 4 (CTLA4) molecules on the T-cell with B7 molecules on the APC. CD28- and CTLA4 molecules bind the same ligand (B7-1 and B7-2 molecules) on the APC, but the CTLA4 molecule has 50-100-fold higher binding activity than the CD28 molecule. The CD28 molecule is expressed by naïve or immature T-cells and also this signal is essential for the initiation and progression of a T-cell response. Interaction

between the CTLA4- and the B7-molecule results in down-regulation of the ongoing immune response. Both the CTLA4 molecule and the inducible co-stimulatory (ICOS) molecule are expressed on activated- and memory T-cells only. ICOS is also important for the T-cell activation and proliferation. It is shown that without ICOS the immune response becomes deficient[12]. Most importantly, ICOS up-regulate the production of interleukin-4 (IL-4), resulting in immunoglobulin-isotype class switching and formation of the germinal centre [13,14].

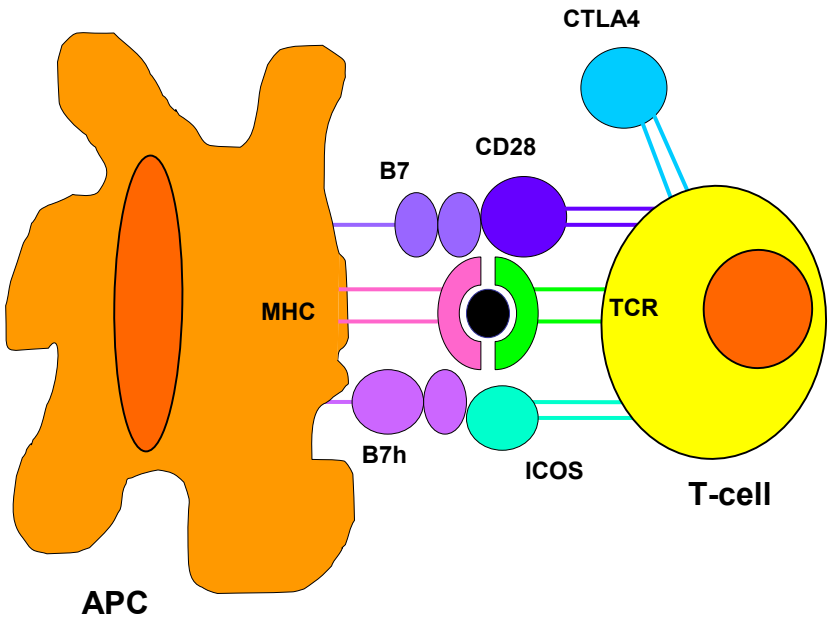

Figure 3: Both the first and second signal between the antigen presenting cell (APC) and the T-cell. The antigen is coloured black and the co-receptors are named in the figure.

4.2.3 The immune response in MS

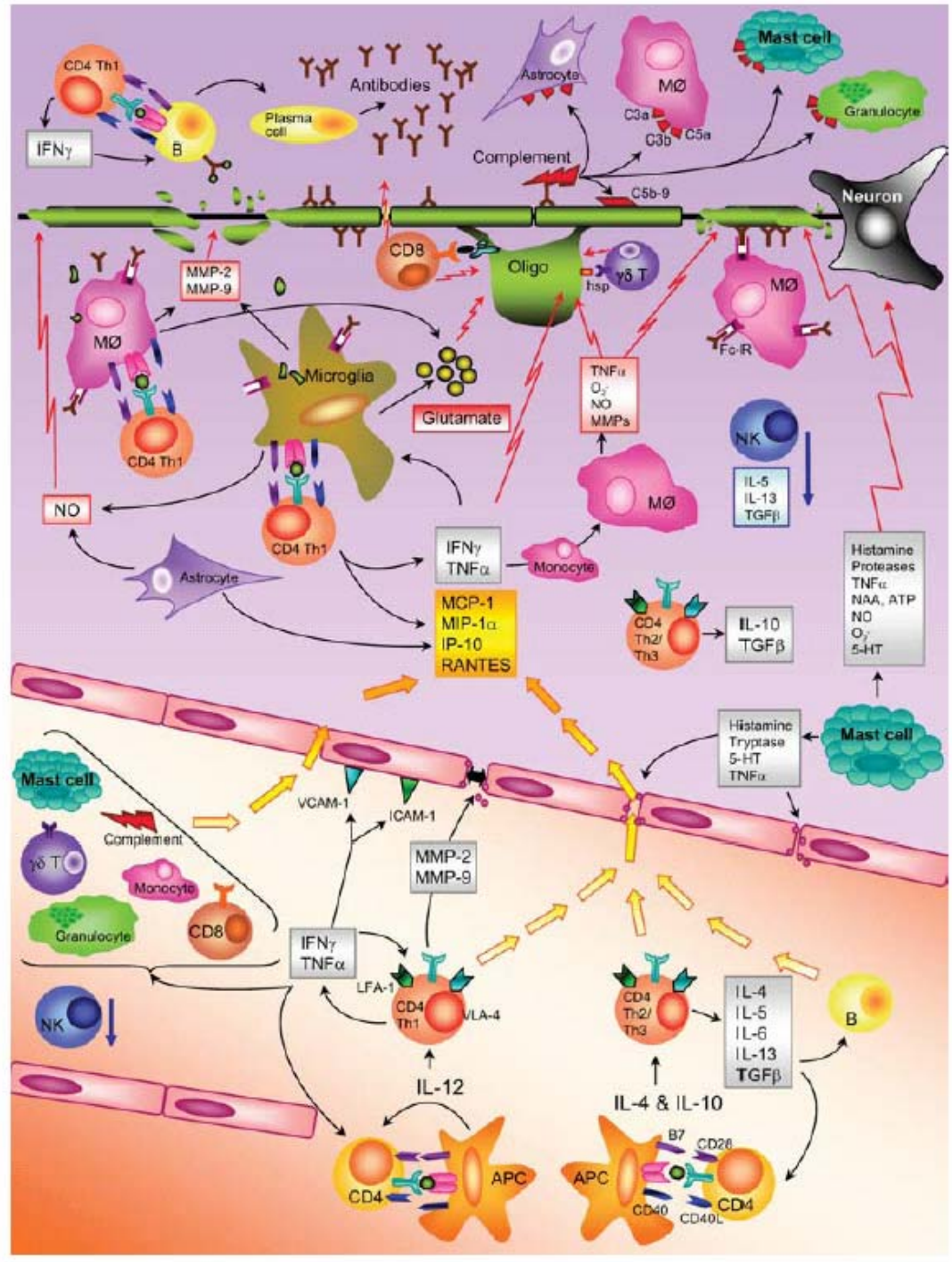

Figure 4: The complex immune response across the blood brain barrier (BBB) of a Multiple Sclerosis (MS) patient. From Sospedra et al, 2004

The immune response in MS is complicated and not fully understood (see Figure 4). The immunological response in CNS is probably characterized by rupture of the blood brain

barrier (BBB) and mononuclear cell infiltration, resulting in destruction of myelin, death of oligodendrocytes and loss of axons. The lesions may occur anywhere in the CNS and vary in size. Most evidence support an autoimmune pathogenesis where lymphocytes, both CD4+ and CD8+ have a fundamental role. Peripherally activated autoreactive CD4+ lymphocytes cross the BBB and initiate chronic inflammatory response in the CNS, as documented in the experimental autoimmune encephalomyelitis (EAE), an animal model of MS[15]. These cells are probably also important players in the long-term evolution of the disease. Damage of the CNS is, however, most likely mediated by other components of the immune system, such as antibodies, complement, $CD8^+$ T cells, and factors produced by innate immune cells. Several mechanisms in MS are still unclear, for example including what guides autoreactive $CD4^+$ T cells to the CNS. However, antigen presentation is a critical requirement for the initiation and perpetuation of inflammatory responses within the CNS[16]. The traditional APCs are not present in the CNS, but MHC class II and costimulatory molecules as B7-1 and B7-2, are upregulated on microglia's and macrophages in the setting of local inflammation. These can effectively present antigens[17].

4.3 Genetic aspects of MS

4.3.1 Genetic epidemiology of MS

About 15-20 % of the MS patients have one or more affected relatives. Family aggregation is often evaluated by studying the relatives of the affected patient and establishing whether they are at higher risk of getting MS than the normal population. Studies have shown that first, second, and third degree relatives of people with MS have a higher risk of developing MS than the general population[18]. But this risk does not differentiate between shared environment and genetic background. Monozygotic twins (MZ) are by birth genetically identical, and by comparing the concordance rates between MZ and dizygotic twins (DZ) the estimation of the role of genetic factors involved in the disease has been done. Figure 5 show the different recurrence risk for a sibling to a MS patient. The risk for getting MS is highest for a monozygotic twin, approximately 25%[19-21] or children of parents who both have MS, approximately 30 %[22,23]. These studies indicate that genes contribute to MS in addition to environmental factors.

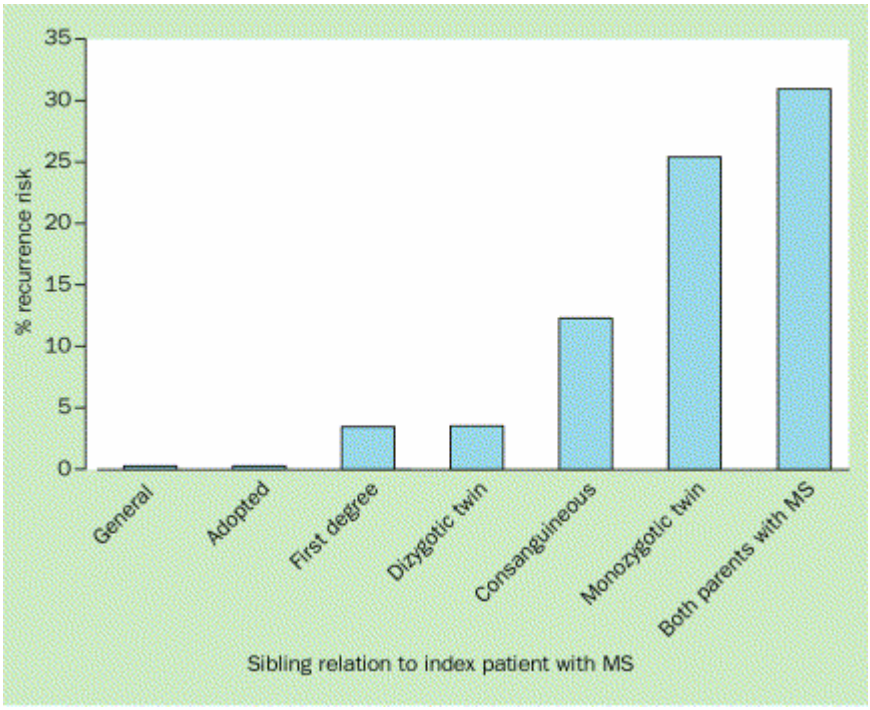

Figure 5: The recurrence risk for getting Multiple Sclerosis (MS) for a sibling to a MS patient. From Dyment et al, 2004

4.3.1 Strategies for genetic studies

Natural sequence variations in the genome can be used as genetic markers. The two main markers used for genetic analysis are microsatellite or variant number of tandem repeats (VNTR) polymorphisms and single nucleotide polymorphism (SNP). **Polymorphisms** are variations in the genome sequence. A locus is polymorphic if there exists two or more normal alleles, with the rarest allele exceeding a frequency of 1 %. Alleles with frequency below 1% are called mutations.

Microsatellites are short, tandemly repeated DNA sequences (Figure 6). The short repeat unit can be di-, tri-, - tetranucleotides or more, but the dinucleotide repeat is the most frequent seen in the genome. The microsatellites are widely spread in the genome and are also highly polymorphic. The mutation rate can be up to 10^{-3} per site per generation compared to an average of 10^{-8} for SNPs.

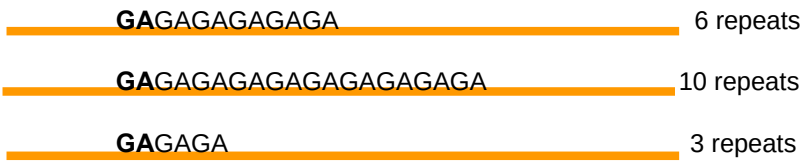

Figure 6: Three DNA segments showing different length of dinucleotide repeats (VNTR).

SNPs are single base pair (bp) variations between normal individuals (Figure 7). The effect of a SNP depends on it's location in the genome, e.g. in the coding region or non-coding region, whether the mutation leads to different incorporated amino acid in the protein synthesis or stop in the protein synthesis. Normally the SNPs are a part of the human variation, but they can also be disease causing.

Figure 7: A single nucleotide polymorphism (SNP) is a point mutation in the DNA sequence.

Genetic analysis is commonly divided into “Linkage analysis” and “Association analysis”. **Linkage** is a term used to describe that two loci on the same chromosome have a tendency to be inherited together more often than by chance alone. Linkage analysis can be used in families to identify genes associated with a specific phenotype. In complex diseases non-parametric sib pair methods referred as a*llele-sharing-test* are used. Due to recombination during the meiosis, most genes are in linkage equilibrium, which means that they are transmitted to the offspring independently of each other. Alleles at loci that are in ***linkage disequilibrium (LD)*** will be transmitted together to the offspring more often than expected from their respective allele frequencies. Alleles in LD tend then to be inherited as so-called haplotype. The recombination rates and the LD throughout the genome will vary, but the LD is always stronger when the loci are close to each other on the chromosome. Regions with high LD are often named conservative regions.
Figure 8 below show the difference between a genotype and a haplotype. A genotype is the alleles specific for a locus, whereas a haplotype is alleles at different loci that are inherited together on the same DNA strand.

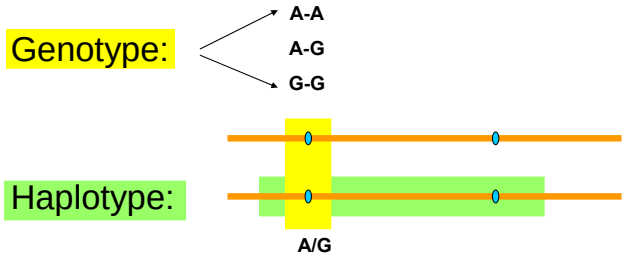

Figure 8: The difference between a genotype and a haplotype.

Association studies test whether a particular allele occurs at higher frequency among affected than unaffected individuals. C*ase-controls studies* are the most common study design in the association studies. In a specific locus the case-controls studies compare the allele frequencies of a set of unrelated affected individuals (the cases) to a set of unrelated unaffected controls (the controls). Another study design is the *family-based association study*. A classical family tree consist of at least three generations, but more

often used for complex disease research is the “TRIO-families”, which consist of the affected persons and theirs parents. The advantage with TRIO based association studies are that the ethnically control material is well matched, since the parent’s non-transmitted alleles are used as controls.

Genetic studies can also be performed as genome-wide screens or candidate gene studies. The **genome wide screens** have mainly been performed using linkage analysis, and have been successfully applied in many monogenetic studies. In studies of the complex traits, as MS, the screens have been more disappointing, and the linkage approach has been complemented with genome wide screens using association analysis. **Candidate gene studies** investigate genes that are selected based on their potential biological relevance to a disease.

4.3.2 Genetic studies in MS

Both candidate gene studies and genome-wide screens have been used to detect genes involved in developing MS. Strongest association and linkage is found to the *HLA* gene region, but several other genes have shown positive findings in many studies, among these are the *T-cell-receptor* α, *SH2D2A* and *CTLA4* genes, and regions at chromosome 5 and chromosome 17.

HLA gene region

The human leukocyte antigen (*HLA*) gene region was the first described gene region, that showed an association with MS and other autoimmune diseases. The *HLA* region involves 3600k bp on the short arm of chromosome 6 (6 p21-22), a highly polymorphic region, which consists of several genes, divided into classes; HLA class I genes (*A*, *B* and *C* loci), *HLA* class II genes (*DR*, *DQ* and *DP* loci) and HLA class III genes (for example *TNF*) (Figure 9)

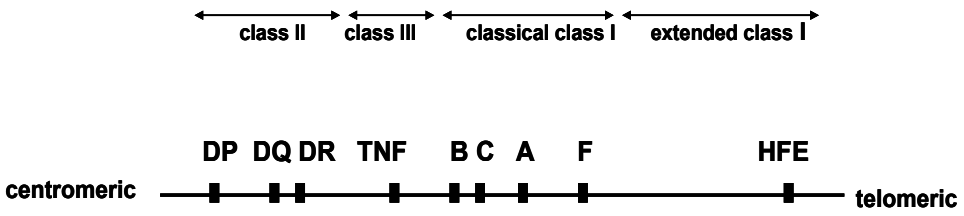

Figure 9: Chromosome 6 and the genes at 6p21-22 in the human leukocyte antigen (HLA) gene region.

The first genetic association with MS was reported in 1972 for *HLA* class I antigens *A3* and *B7*[24,25]. Some years later association with MS to *HLA* class II was shown. The strongest and most consistent association is found with the extended haplotype *HLA DRB1*1501*, *DQA1*0102*, *DQB1*0602 (DR1501 or DR2)* haplotype[26]. This haplotype is especially frequent in Northern Europe. In a Norwegian material of 243 MS patients and 296 randomly colleted Norwegian controls, the frequency for the *HLA DRB1*1501*, *DQA1*0102*, *DQB1*0602 haplotype* was found to be 36.2% among MS patients and

16.5% among the controls[27]. Similar frequencies are also found in British and Swedish studies[26,28]. The *HLA-DR2* haplotype was more frequent among female patients, and was negatively correlated with age at time of diagnosis regardless of sex both in an English, Swedish and Norwegian study[1,28,29].

Several non-northern European populations have different *DRB1* associations with MS. Both MS in Sardinia, Canary Islands and Turkey have in additional to the *HLA DRB1 15* haplotype an association with the *DR4* allele (*DRB1*0405*)[30-32].

It has also been reported that genes located in the *HLA* complex outside the class II region are associated with MS. Among Swedish patients, the *HLA-A*0301* allele was found to increase the risk of MS independently of the *HLA-DR2* haplotype, whereas the *HLA-A*0201* allele was found to decrease the risk of MS[33]. Similar results were recently replicated in a Norwegian study, where we found an independent additional effect to genetic susceptibility of HLA class I in MS. The study found a higher "genotypic relative risk" in individuals carrying both *HLA DR2* and *HLA A3 (HLA-A*0301)*, compared to those who carry only *HLA DR2* or only *HLA A3*[27].

Genome-wide Screens for Linkage in MS

The first three genome-wide screens for linkage studies preformed in MS were published in 1996[34-36]. None of these screens could define new susceptibility loci, but there were more regions with suggestive and potential linkage, than expected by chance alone. Since these first publications, several genome-wide screens have been performed in other populations, including the Nordic countries[37,38]. However, few screens have overlapping results. In 2004 a large meta-analysis with 719 families from the Europe, USA and Canada was preformed. This screen reveled for the first time significant linkage in the HLA region in MS due to improved power caused by including a high number of families [39].

GAMES (Genetic Analysis of Multiple sclerosis in EuropeanS) was a big European collaboration project performing a genome-wide screen for LD in MS. By using a collaborative network, large datasets of cases and controls were collected. The individual DNA samples were pooled together and around 6000 microsatellite markers were typed

for these pools of DNA. Follow-up studies of the GAMES screens are in progress[40].

Other candidate gene regions in MS

A long list of candidate genes has been studied in MS. Among the most interesting genes is the *CTLA4* gene, which is the main focus at this present work (see next paragraph). In addition, the SH2 domain protein 2A gene (***SH2D2A*)**, located in the chromosome 1q21 region, has been an interesting candidate gene to study (see related paper[41,42]). This gene encodes the T cell specific adaptor protein (TSAd), which was first described to be expressed in activated human T-cells[43]. It has been implicated in modulation of proximal signaling events, as well as in transcriptional regulation in human T-cells. A variant number of tandem repeats (VNTR) polymorphism was found in the gene's promoter region. Recent studies indicate that homozygosity for short (ie GA(13) and GA(16)) alleles for this polymorphism, is associated with development of MS[41]. This finding is replicated in other autoimmune diseases such as juvenile rheumatoid arthritis (JRA)[42]. Thus, these studies support the notion that the short alleles of the *SH2D2A* promoter polymorphism may contribute to the genetic susceptibility in autoimmune diseases.

4.3.3 The *CD28/CTLA4/ICOS* gene region

Chromosome 2q33 encodes a set of genes, which are essential for several functions in the T-cell. Among these are the *CTLA4* gene, the *CD28* gene and the *ICOS* gene (Figure 10).

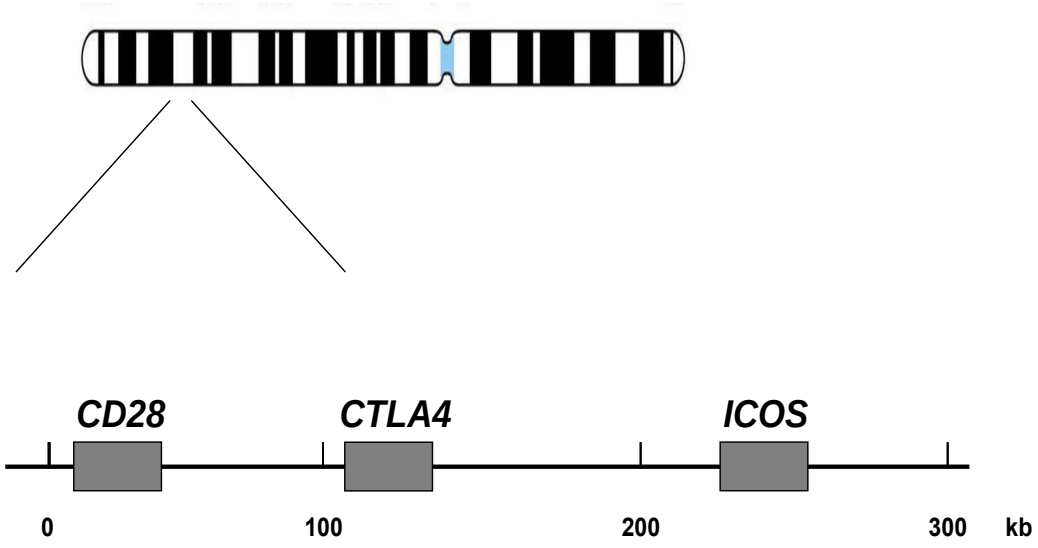

Figure 10: Chromosome 2 and the genes at 2q33. From centromeric the cluster of differentiation 28 (*CD28*) gene, the cytotoxic T lymphocyte antigen 4 (*CTLA4*) and most telomeric the inducible co-stimulatory (*ICOS*) gene.

These three molecules are all located on the T-cell surface showed in Figure 11. These genes have been of interest since they all play an important role in the immune response. The hypothesis is that mutations or polymorphims in these genes resulting in a dysfunction in the molecules can implicate and disrupt the immune response and finally contribute to an autoimmune disease.

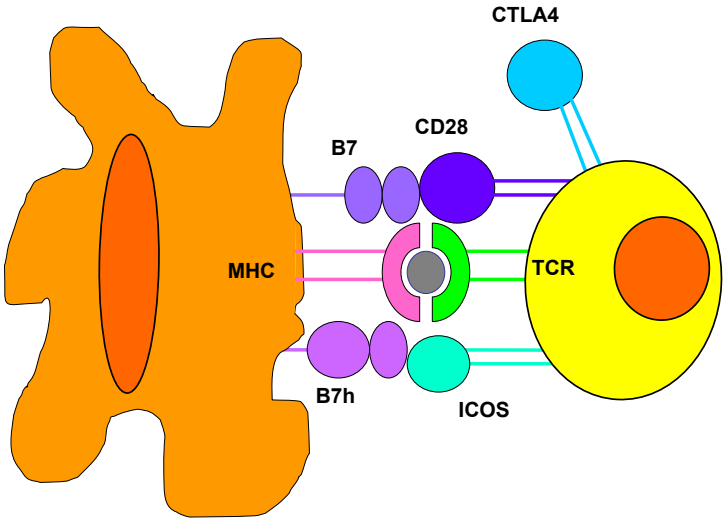

Figure 11: The different cell surface molecules on the antigen presenting cell (APC) and the T-cell. The antigen is coloured grey and the co-receptors are all named in the figure.

Especially, the *CTLA4* gene has been shown to be linked to and associated with several autoimmune diseases[44]. In a Norwegian material an association with the *CTLA4+49* polymorphism in MS was reported in 1999. A series of genetic analyses of the *CTLA4* gene and other closely located genes have thereafter been reported in MS (for overview see Teutsch et al[45]), but diverse findings have been reported. An increase in disability has been found to be associated with the *CTLA4 +49* AA genotype[46], as well as association of the G allele with a PP-MS disease course[47]. Other studies have not shown associations neither with age at onset, disease course nor severity[48,49].

Recently, a new candidate gene polymorphism in the *CTLA4* region was reported to influence gene splicing and thereby the relative abundance of soluble CTLA4[50]. A decrease in soluble CTLA4 protein will result in incomplete down-regulating of the immune response, a mechanism that may be involved in development of autoimmune disease. Association was shown for a haplotype including the *CT60* single nucleotide polymorphism (SNP) (located 0.2 kb 3' of the *CTLA4* transcript), in Graves' disease, autoimmune hypothyroidism and type 1 diabetes[50]. Association with this gene region has

also been found in celiac disease[51] and Addison's disease[52]. Few studies have shown an association with autoimmune diseases to the *CD28* and *ICOS* genes in spite of the important functions of these genes.

On this background, this work has focused on a candidate gene association study in the *CD28/CTLA4/ICOS* gene region in MS.

5 AIMS OF THE STUDY

The overall aim of this study has been to investigate genetic factors contributing to disease susceptibility to MS.

The specific aims of this work have been to:

1) Reevaluate the previous association found in the *CTLA4* region with MS in a new material. The investigated region has been expanded and additional markers have been used to find out if the previously association is primary or secondary to a nearby region.
2) Analyze distribution of alleles in different clinical and genetic subgroups of MS.

6 METHODOLOGICAL CONSIDERATIONS

6.1 The patients and controls

Most of the MS patients included in this study are collected in Oslo or its suburban areas. The majority of these patients were recruited through their contacts with Department of Neurology, Ullevål University Hospital. A few were recruited through MS Societies and other neurological departments serving the suburban Oslo areas. An additional population of MS patients was recruited from the Department of Neurology, Haukeland University Hospital in Bergen. All patients included in this study have been diagnosed in neurological departments, and all fulfilled the Poser criteria for MS[6].

All control samples were randomly collected among healthy blood donors recruited through the Norwegian Bone Marrow Registry. The study was approved by the Regional Medical Research Ethics Committee.

6.2 Genotyping

6.2.1 The PCR method

The first publication of the polymerase chain reaction (PCR) technique was released in 1985[53]. The technique amplifies a specific DNA segment, and the method is sensitive enough to detect as little as one DNA molecule in almost any type of sample, as blood-, hair root- and sputum/mucous membrane samples.

Selective amplification requires information about the DNA sequences flanking the target DNA. This information can be found in genetic databases or detected through sequencing the region. Based on this information, two oligonucleotides or primers are designed. The primers are complementary to short sequences outside the 3'endes on the DNA strands. PCR is a chain reaction where the synthesized DNA strands act as template for further DNA synthesis. A PCR-cycle consists of at least three steps, and the steps are temperature- and time specific. The first step is the denaturation of the double stranded DNA, the second step is the annealing of the primers. The third step is the DNA synthesis using a heat-stable DNA polymerase (Taq Polymerase).

The number of cycles are usually between 25 and 30, resulting in a specific DNA segment in at least 10^5 copies. The specific DNA segment can be visualized as a distinct

band of a specific size on a gel electrophoresis.

6.2.2 The PCRs in this study

The primers were either obtained through the Genome Database (http://ww.gdb.org) or established in previous projects. The *CT60* (rs3087243) and *CT61 (*rs11571319) SNPs were genotyped as one haplotype by amplifying a polymerase chain reaction (PCR) product including both SNPs, which was separated by 18 base pairs (forward primer 5'TATCCATCCTCTTTCCTTTTGA, reverse primer 5'AAATCAATTGGCATGCTGTTTAAC).

The microsatellite markers were amplified separately by PCR in a reaction volume of 10 µl containing low salt buffer, 8 nmol dNTP, 1.6 pmol of each primer, 0.16 U Taq Polymerase and 40 ng DNA or in a reaction volume of 8 µl containing 1xMastermix (http://www.ahdiagnostics.com), 2.2 pmol of each primer and 40 ng of DNA.

The SNP markers were amplified in a 20 µl reaction volume using $MgCl_2$-free PCR buffer, 50 nmol $MgCl_2$, 8 nmol dNTP, 4 pmol of each primer, 0.8 U Taq Polymerase and 20ng of DNA.

Thermal cycling conditions were: 95°C 20 seconds, 56-60°C 30 seconds (depending on the annealing temperature), 72°C 60 seconds, 35 cycles ran on a Peltier Thermal Cycler (MJR Research, Watertown, MA).

6.2.3 Genotyping by use of PCR restriction length polymorphism method

Restriction enzymes are enzymes able to cut DNA into fragments. A given enzyme recognizes a specific sequence of 4-8 nucleotides, and this is called the restriction site. The PCR-restriction length polymorphism (PCR-RFLP) method is based on a difference in one nucleotide among individuals, and is therefore a method for detecting SNPs. The PCR product has the SNP inside the fragment and the SNP is also located in a restriction site for a given restriction enzyme. Depending on the nucleotide in the SNP position the used restriction enzyme will or will not recognize the site and either cut or not cut the DNA segment. There are three possible genotypes for each individual sample; homozygote for wild or mutant nucleotide in the SNP, or heterozygote (see Figure 12).

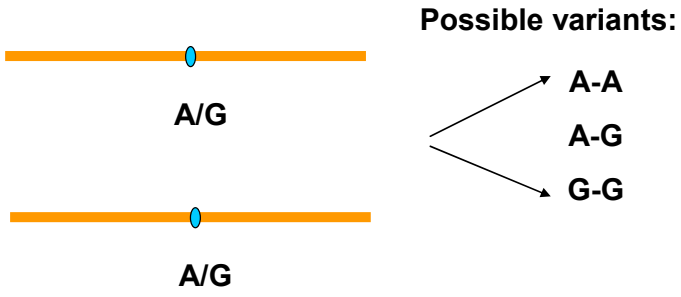

Figure 12: There are three possible genotypes at a polymorphic locus.

The DNA fragment from the individual samples will be visualized as distinct bands on a gel electrophoresis. With standard bands in the same run, the genotype for each individual sample can be determined.

6.2.4 Genotyping by use of TaqMan® technology

TaqMan® Technology integrates a PCR-based assay with laser scanning technology to excite fluorescent dyes present in the specially designed TaqMan® probes. The ABI PRISM® 7000 Sequence Detection System (Applied Biosystems, Foster city, California, USA) can be used for assays such as allelic discrimination (SNP detection). In our study we used a 384-well assay. Data were collected and analyzed with software from Applied Biosystems, TaqMan® program SDS2.1. Negative controls and controls with known genotype for the marker (homozygote or heterozygote) were run in every PCR subset.

6.2.5 Genotyping by use of DNA sequencer

The PCR products can be amplified using primers labeled with different fluorochromes. The length of the PCR product can be detected when these PCR products are separated in a DNA sequencer together with an internal size standard. In our study the microsatellite alleles were identified according to length by analyzing the PCR products on an ABI Prism® XL 337 sequencer (Applied Biosystems, Foster city, California, USA). Data were collected and analyzed with the Gene Scan® 2.1 and the Genotyper® 2.0 programs. Controls with known length were included in each run, and blank control samples were

run in every PCR subset.

6.2.6 Genotyping by use of melting gel electrophoresis

The melting gel electrophoresis theory is based on the observation that double-stranded DNA fragments melt to single-stranded DNA when exposed to denaturants such as urea, formamide and/or temperature. The double-stranded DNA fragments have different melting properties depending on the nucleotide sequence and length of the fragments. When exposed to denaturants the PCR amplified fragments with only minor difference like a point mutation, can be detected based on their differential melting point and migration in a polymer-filled capillary. *Fischer SG et al*[54] have described the methodology in more details.

Denaturant capillary electrophoresis (DCE) is based on the melting gel theory[55-57], and is a method for genotyping of SNPs. By comparing the peak pattern with a heterozygous standard run simultaneously with a different florochrome, the individual samples could be genotyped. The SNPs *CTLA4+49, CT60* and *CT61* were genotyped using DCE.

The SNP typing was performed on a standard multicapillary DNA sequencing instrument MegaBASE™ 1000 DNA Analysis System (Amersham Pharmacia Biotech, Oslo, Norway). The *CTLA4 +49* SNP was separated by a gradient of 58-54°C cycled 20 times, followed by constant temperature at 50°C. The *CT60* and *CT61* alleles (i.e. the combination of the *CT60* and *CT6*1 microhaplotypes) were separated by cycling the temperature between 50-47°C 40 times, followed by lowering the temperature to 45°C until the end of the run. The peak patterns were then analyzed and compared with a heterozygous standard by a Genotype program named “View and Edit” from MegaBASE™ software.

6.3 Statistical considerations

In our study the case and control materials were compared by the Chi square test using the Public Domain Software for Epidemiology and Disease Surveillance EPI Info Version 5.01b (Center of Disease Control, Epidemiology Program Office, Atlanta, GA, USA). Global chi square values and global p-values were calculated for each marker, after having grouped alleles with frequencies below 0.05 in cases and controls into one group. Since none of the global p-values reached statistical significance at the level of 0.05, only uncorrected p-values are given. The haplotype frequencies were estimated by the expectation-maximization (EM) algorithm using the COCAPHASE program available through the Genetic Linkage User Environment (GLUE) application at Human Genome Mapping Project Resource center (HGMP-RC) (www.hgmp.mrc.ac.uk). Hardy-Weinberg equilibrium was calculated manually or using Arlequin software (http://anthro.unige.ch/arlequin). Clinical correlations and logistic regression analysis was performed using NCSS 2004 (Number Cruncher Statistical Systems; www.ncss.com).

7 SUMMARY OF RESULTS

We have investigated in this study a 262 kb region of chromosome 2q33, covering the *CD28, ICOS* and *CTLA4* genes, in a clinically well characterized Norwegian MS cohort. Six microsatellites; *CD28-A, CD28-B, SARA-43, SARA-1, SARA-31* and *SARA-47* and three SNPs; *CTLA4+49*, *CT60* and *CT61* were genotyped (Figure 13).

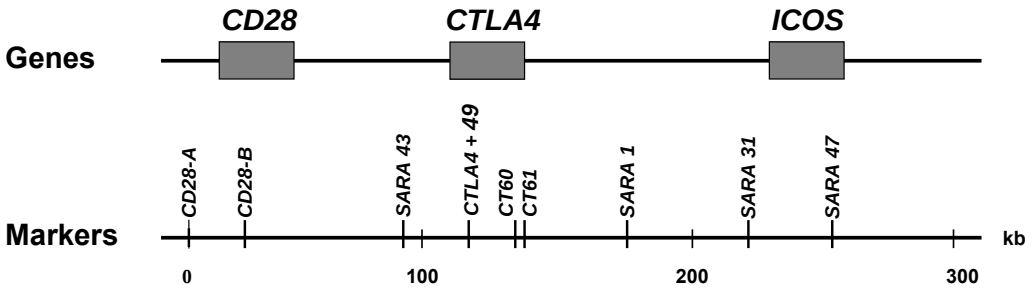

Figure 13: Genes and markers at chromosome 2q33 investigated in this study.

All result tables are placed at the end of this chapter.

DCE methodology was selected for the SNP genotyping-

When starting up the present study, we experienced some technical difficulties using the PCR-RFLP methodology that was applied in our previous study[58]. In that study the amplified product was determined by sequential digestion with the restriction enzymes TseI and MseI. We performed an explorative comparison between the different SNP genotyping methods; PCR-RFLP-, TaqMan®- and denaturant capillary electrophoresis (DCE)-analysis.

Genotyping of the *CTLA4+49* marker in 189 controls showed different results (Table 3) in 25 samples when comparing PCR-RFLP data with TaqMan® data, and in 24 samples when comparing PCR-RFLP data and DCE data (typing error 13 %). When comparing TaqMan® and the DCE data, two samples were genotyped with different result (typing error 1 %).

Table 4 shows the comparison between the different methods when analyzing the specific genotypes. In the RFLP-PCR method the GG-genotype was more often found, than using other methods as TaqMan® and DCE. Based on these results, the DCE method was selected for further SNP typing, and only the DCE data are included in the analyses presented in this study.

No associations were found to markers in the CD28/CTLA4/ICOS gene region

Genotyping of the nine markers in the *CD28/CTLA4/ICOS* gene region (*CD28-A, CD28-B, SARA-43, CTLA4+49, CT60*, *CT61, SARA-1, SARA-31* and *SARA-47*) in the first material of MS patients (n=302) and controls (n=282), showed no significant differences in the global analysis of allele frequencies (Table 5).

Since the *CTLA4* gene previously has been reported to be associated with MS in some studies, a second material of Norwegian MS cases (n=273) and controls (n=269) was genotyped for the *CTLA4, CT60* and *CT61* SNPs. No significant deviations in allele frequencies or genotype frequencies were found in this second material (Table 6). The association was neither not found in the total material of 575 MS cases and 551 controls (Table 6).

CTLA4+49, CT60, *CT61* haplotype frequencies in the MS cases and controls were also compared without showing significant associations (Table 7).

No associations were found after stratification of the material

Stratifying the part of the material that previously has been typed for the HLA-DR1502, DQ0602 haplotype (n= 274 MS cases and n=300 controls), did not reveal additional information (data not shown).

Furthermore, the frequencies of genotypes and haplotypes for markers in the *CTLA4* region showed no correlations neither to sex (n= 485), disease course (n= 453), nor

presence or absence of familial cases of MS (n= 448) (Table 8).There were no associations to age at onset (data not shown). Logistic regression analysis including these clinical variables and genotyping data did not reveal additional information.

Power and Hardy-Weinberg calculations

All SNP markers in the control material were found to be in Hardy Weinberg equilibrium (data not shown). Power calculations for the six microsatellites were done, showing that this study could not show an OR below 1.5.

Result tables in this study:

	PCR-RFLP	TaqMan®	DCE
Genotype			
AA	0.32	0.33	0.33
AG	0.46	0.49	0.49
GG	0.22	0.18	0.18
Allele			
A	0.55	0.57	0.57
G	0.45	0.43	0.43

Table 3: Genotype and allele frequencies of the *CTLA4+49* polymorphism in 189 randomly collected Norwegian healthy controls achieved by different methods.

	PCR-RFLP to TaqMan®	PCR-PFLP to DCE	TaqMan® to DCE
Differents			
GG > GA	16	16	1
GA > GG	8	7	1
GA > AA	1	1	
	25	24	2
Typing error freqency	0.13	0.13	0.01

Table 4: Comparing genotyping errors using different methods.

Allele	MS	Control	[a] x^2_{global}	[b] $p_{global\ nc}$
[c] ***CD28-A***	**2n=594**	**2n=564**	4.30	0.25
231	0.06	0.03		
235	0.81	0.85		
237	0.07	0.05		
other	0.07	0.07		
[c] ***CD28-B***	**2n=590**	**2n=560**	5.57	0.23
198	0.33	0.39		
200	0.18	0.18		
202	0.12	0.19		
204	0.35	0.39		
other	0.01	0.01		
[c] ***SARA-43***	**2n=592**	**2n=498**	10.24	0.07
218	0.14	0.11		
220	0.09	0.06		
222	0.49	0.55		
224	0.07	0.08		
226	0.16	0.13		
other	0.05	0.06		
[d] ***CTLA4 +49***	**n=240**	**n=241**	0.71	0.40
A	0.55	0.59		
G	0.45	0.41		
[d] ***CT60***	**n=261**	**n=267**	0.54	0.46
A	0.40	0.43		
G	0.60	0.57		
[d] ***CT61***	**n=261**	**n=267**	0.35	0.56
A	0.15	0.17		
G	0.85	0.83		
[c] ***SARA-1***	**2n=594**	**2n=498**	3.27	0.35
260	0.05	0.06		
264	0.29	0.33		
268	0.28	0.27		
272	0.36	0.32		
other	0.02	0.02		
[c] ***SARA-31***	**2n=602**	**2n=498**	1.69	0.43
211	0.50	0.51		
217	0.49	0.49		
other	0.00	0.00		
[c] ***SARA-47***	**2n=604**	**2n=498**	2.86	0.41
149	0.37	0.33		
151	0.34	0.34		
153	0.24	0.27		
other	0.05	0.06		

Table 5: Allele frequencies for analyzed chromosome 2q33 markers in the first Norwegian MS patients and controls material.

[a] x^2_{global} = global chi square value, not corrected for number of comparisons

[b] $p_{global\ nc}$ = global p-value, not corrected for number of comparisons

Genotype	[a] MS I	[b] MS II	[a] Control I	[b] Control II	MS_{total}	$Control_{total}$	[c] $x^2_{global\ total}$	[d] $p_{global\ nc\ total}$
CTLA4 +49	**n=240**	**n= 273**	**n=241**	**n=268**	**n=513**	**n=509**	0.37	0.83
AA	0.29	0.34	0.35	0.32	0.32	0.33		
AG	0.52	0.47	0.49	0.49	0.49	0.49		
GG	0.19	0.19	0.16	0.19	0.19	0.18		
CT60	**n=261**	**n=217**	**n=267**	**n=269**	**n=478**	**n=536**	0.59	0.75
AA	0.17	0.18	0.18	0.15	0.18	0.17		
AG	0.45	0.49	0.49	0.50	0.47	0.49		
GG	0.38	0.33	0.33	0.35	0.35	0.34		
CT61	**n=261**	**n=217**	**n=267**	**n=269**	**n=478**	**n=536**	0.48	0.48
AA	0.02	0.02	0.04	0.01	0.02	0.03		
AG	0.27	0.26	0.25	0.29	0.27	0.27		
GG	0.71	0.72	0.71	0.70	0.71	0.70		
Allel	[a] MS I	[b] MS II	[a] Control I	[b] Control II	MS_{total}	$Control_{total}$	[c] $x^2_{global\ total}$	[d] $p_{global\ nc\ total}$
CTLA4 +49	**n=240**	**n= 273**	**n=241**	**n=268**	**2n=1026**	**2n=1018**	0.37	0.54
A	0.55	0.58	0.59	0.57	0.57	0.58		
G	0.45	0.42	0.41	0.43	0.43	0.42		
CT60	**n=261**	**n=217**	**n=267**	**n=269**	**2n=956**	**2n=1072**	0.01	0.91
A	0.40	0.43	0.43	0.40	0.41	0.41		
G	0.60	0.57	0.57	0.60	0.59	0.59		
CT61	**n=261**	**n=217**	**n=267**	**n=269**	**2n=956**	**2n=1072**	0.51	0.48
A	0.15	0.15	0.17	0.16	0.15	0.16		
G	0.85	0.85	0.83	0.84	0.85	0.84		

Table 6: *CTLA4 +49*, *CT60* and *CT61* allele and genotype frequencies in the different sets of Norwegian MS patients and controls.

[a] the first material of Norwegian multiple sclerosis (MS) cases and randomly collected healthy controls

[b] the second material of Norwegian MS cases and controls

[c] $x^2_{global\ all}$ = global chi square value for the total material

[d] $p_{global\ nc\ all}$ = global p-value for the total material, not corrected for number of comparisons

Haplotype	[a] MS	[a] Control	[b] x^2_{global}	[c] $p_{global\ nc}$
	2n=878	2n=974		
A, A, G	0.40	0.40	0.04	0.85
A, G, A	0.14	0.16	0.51	0.47
G, G, G	0.44	0.42	0.38	0.54
other	0.02	0,02		

Table 7: *CTLA4+49*, *CT60*, *CT61* haplotype frequencies in Norwegian MS patients and controls.

[a] total population of multiple sclerosis (MS) patients and controls

[b] x^2global = global chi square value

[c] $p_{global\ nc}$ = global p-value, not corrected for number of comparisons

	Sex		Disease course		Familial cases of MS	
	female	male	[a] RRMS	[b] PPMS	no	yes
Genotype						
CTLA4+49	n=334	n=151	n=372	n=81	n=373	n=75
AA	0.32	0.30	0.34	0.25	0.32	0.28
AG	0.50	0.50	0.48	0.50	0.49	0.49
GG	0.18	0.20	0.18	0.25	0.19	0.23
CT60	n=326	n=145	n=363	n=82	n=360	n=77
AA	0.18	0.15	0.19	0.15	0.18	0.18
AG	0.46	0.46	0.47	0.41	0.40	0.47
GG	0.34	0.39	0.34	0.44	0.42	0.35
CT61	n=326	n=145	n=363	n=82	n=360	n=77
AA	0.01	0.03	0.01	0.03	0.00	0.02
AG	0.26	0.27	0.28	0.21	0.25	0.27
GG	0.73	0.70	0.71	0.76	0.75	0.71
Allele						
CTLA4+49	2n=666	2n=302	2n=742	2n=162	2n=744	2n=150
A	0.57	0.55	0.58	0.50	0.53	0.57
G	0.43	0.45	0.42	0.50	0.47	0.43
CT60	2n=652	2n=290	2n=726	2n=164	2n=720	2n=154
A	0.42	0.38	0.43	0.35	0.38	0.41
G	0.58	0.62	0.57	0.65	0.62	0.59
CT61	2n=652	2n=290	2n=726	2n=164	2n=720	2n=154
A	0.14	0.16	0.15	0.14	0.12	0.15
G	0.86	0.84	0.85	0.86	0.88	0.85

Table 8: Allele and genotype frequencies in clinical subgroups according to disease course and demographic variables in Norwegian MS patients.

a RRMS = relapsing remitting disease course

b PPMS = primary progressive disease course

8 GENERAL DISCUSSION

Chromosome region 2q33 encodes several regulators of the immune system, among these the CTLA4, CD28 and ICOS molecules. Involvement of these genes in MS is not yet clear. We investigated six microsatellites and three SNPs in a relatively large and clinically well characterised Norwegian MS cohort. No associations were observed for any of the markers analysed in 575 MS patients and 551 controls. Associations were neither found when stratifying the material for the HLA-DR1501 haplotype, sex, age at onset, disease course nor familial aggregation. In conclusion, this study could not confirm association with the *CD28/CTLA4/ICOS* gene region.

When investigating candidate genes and genetic markers, the initial positive report of an association, is often followed by subsequent negative findings. A recent review reported that up to 95% of initial claims of association cannot be replicated[59]. The authors suggested that the most important factors behind the inability to replicate previous positive association were "publication bias, failure to attribute results to chance and inadequate samples sizes". Several of these factors may have contributed to the fact that this study can not confirm previously positive findings.

Our methological evaluation identified a divergent result between the PCR-RFLP method and the two other methods (DCE and the TaqMan®) (Table 3 and 4). These analyses showed that both the DCE and the TaqMan® methods were superior to the RFLP method for accurate genotyping of the selected SNPs. Incorrect genotype frequencies might therefore have appeared in both our own and other's previous studies reporting RFLP based *CTLA4+49* analyses. Methodical evaluations are seldom reported, but are obviously of great value to a genetic study.

All available MS patients and controls were genotyped for the *CTLA4*, *CT60* and *CT61* SNPs by using the DCE methodology in the present study. DCE was selected before TaqMan® genotyping due to our previous experience with this methodology showing that DCE is a reliable and low cost tool for high-throughput SNP analysis. The previously

obtained RFLP data were only included in methodology studies. In order to present data that are consistently genotyped by use of the same methodology, only DCE SNP data are included in all statistical analyses. Since all the material of MS patients and controls SNPs now only is genotyped by using DCE technology, the frequency of genotyping inconsistencies should be low in both the first and second Norwegian MS cohorts presented in this study.

Non-significant, but interesting differences in genotype frequencies between the two materials were still observed in this study (Table 6). Both materials are established by random selection among Norwegians, thus ethnical differences can not explain the differences between the two sets of Norwegian cases and controls. Even though it is very important to match the control group with the patient group, an undetected stratification may have appeared in the control material. Recently, an Icelandic study showed that the difference between "randomly collected controls" is more divergent than excepted[60]. To reduce this stratification risk one might rather analyze TRIOs (the patients and their parents) instead of case-controls materials. The parents can then be used as controls, and Mendelian consistency can be used as an internal check for genotype errors.

It has been suggested that "inadequate samples sizes" is an important factor behind the inability to replicate previous positive associations[59]. The results obtained in published MS studies about the *CTLA4* region, have been inconsistent[45,61], probably partly due to small datasets. Usally below 300 cases and 300 controls or below 185 families have been included[47-49,58,62-64]. A recent meta-analysis of nine of the Caucasian datasets of MS patients typed for the *CTLA4+49* polymorphism, did not reveal significant associations[45].Table 9 shows this analysis where OR for all the studies together are close to one, showing that there is no significant association with MS in these *CTLA4* studies.

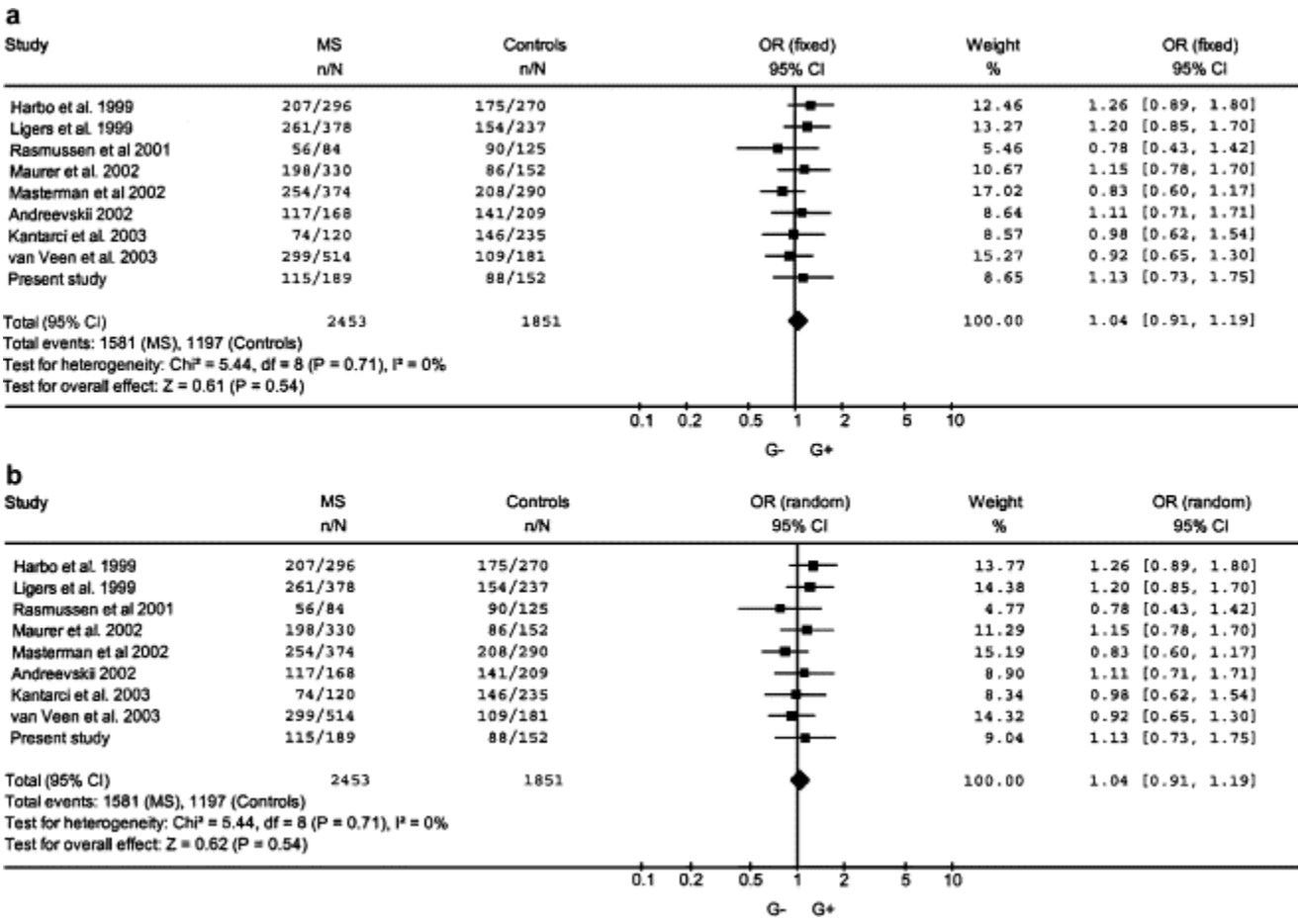

Table 9: Meta-analysis of *CTLA-4* exon 1 +49 G allele carrier frequency in nine comparable datasets of MS patients and controls (a) Data analyzed using a fixed effects model. (b) Data analyzed using a random effects model. From Teutsch et al, 2004

The meta-analysis of seven of the same datasets comparing the *CTLA4+49* polymorphism in MS patients with different clinical courses did not show associations (not shown)[45]. The first Norwegian report[58] was included in these meta-analyses. In line with this meta-analysis, the differences between cases and controls declined when we combined our data sets to a large MS case and control cohort. This observation emphasizes strongly the importance of including large materials in case-control analyses of candidate genes in a complex disease like MS.

Association with the *CTLA4* gene region has been found in many autoimmune diseases, among these are Graves' disease, autoimmune hypothyroidism, type 1 diabetes[50], celiac disease[51] and Addison's disease[52].These findings have been of importance for the postulation that the *CTLA4* gene is important in developing autoimmune diseases. As previously mentioned, studies on gene splicing have found an influence between slicing forms and relative abundance of soluble CTLA4[50]. A decrease in soluble CTLA4 protein resulted in incomplete down-regulating of the immune response, a mechanism that have been postulated to be involved in development of autoimmune diseases. However, it seems like MS is not associated with *CTLA4,* in contrast to some other autoimmune diseases. This could indicate that the pathogenesis of MS differ somehow from the pathogenesis of some other autoimmune diseases.

9 FUTURE STUDIES

The study of the *CD28/CTLA4/ICOS* gene region in MS presented here, adds to the findings in other recent publications that this genetic region dose not seem to make a major contribution to the genetic susceptibility in MS. Thus, our future studies will be focused on other candidate regions, as the *HLA class I* region at chromosome 6 and the *SH2D2A* gene at chromosome 1. Since studies in these genetic regions have already been performed (see paper 2 and 3) with interesting findings, further studies are ongoing.

10 REFERENCES

1 Celius,E.G. and Vandvik,B. (2001) Multiple sclerosis in Oslo, Norway: prevalence on 1 January 1995 and incidence over a 25-year period. *Eur. J. Neurol.* 8, 463-469

2 SWANK,R.L. *et al.* (1952) Multiple sclerosis in rural Norway its geographic and occupational incidence in relation to nutrition. *N. Engl. J. Med.* 246, 722-728

3 Westlund,K. (1970) Distribution and mortality time trend of multiple sclerosis and some other diseases in Norway. *Acta Neurol. Scand.* 46, 455-483

4 Dahl,O.P. *et al.* (2004) Multiple sclerosis in Nord-Trondelag County, Norway: a prevalence and incidence study. *Acta Neurol. Scand.* 109, 378-384

5 Charcot JM (1868) Histologie de la sclérose en plaques. *Gaz. Hop. civils et militaires*

6 Poser,C.M. *et al.* (1983) New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann. Neurol.* 13, 227-231

7 McDonald,W.I. *et al.* (2001) Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann. Neurol.* 50, 121-127

8 KURTZKE,J.F. (1983) Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33, 1444-1452

9 Ramsaransing,G. *et al.* (2001) Early prediction of a benign course of multiple sclerosis on clinical grounds: a systematic review. *Mult. Scler.* 7, 345-347

10 Kay,A.B. (1994) Origin of type 2 helper T cells. *N. Engl. J. Med.* 330, 567-569

11 Robert C.Mellors. Immunopathology. 27-8-2002. Cornell University Medical College.

12 Richter,G. and Burdach,S. (2004) ICOS: a new costimulatory ligand/receptor pair and its role in T-cell activion. *Onkologie.* 27, 91-95

13 Dong,C. *et al.* (2001) ICOS co-stimulatory receptor is essential for T-cell activation and function. *Nature* 409, 97-101

14 Rudd,C.E. and Schneider,H. (2003) Unifying concepts in CD28, ICOS and CTLA4 co-receptor signalling. *Nat. Rev. Immunol.* 3, 544-556

15 Hickey,W.F. *et al.* (1991) T-lymphocyte entry into the central nervous system. *J. Neurosci. Res.* 28, 254-260

16 Dong,Y. and Benveniste,E.N. (2001) Immune function of astrocytes. *Glia* 36, 180-190

17 Markovic-Plese,S. and McFarland,H.F. (2001) Immunopathogenesis of the multiple sclerosis lesion. *Curr. Neurol. Neurosci. Rep.* 1, 257-262

18 Sadovnick,A.D. *et al.* (1988) Multiple sclerosis: updated risks for relatives. *Am. J. Med. Genet.* 29, 533-541

19 Ebers,G.C. *et al.* (1986) A population-based study of multiple sclerosis in twins. *N. Engl. J. Med.* 315, 1638-1642

20 Sadovnick,A.D. *et al.* (1993) A population-based study of multiple sclerosis in twins: update. *Ann. Neurol.* 33, 281-285

21 Willer,C.J. *et al.* (2003) Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc. Natl. Acad. Sci. U. S. A* 100, 12877-12882

22 Ebers,G.C. *et al.* (2000) Conjugal multiple sclerosis: population-based prevalence and recurrence risks in offspring. Canadian Collaborative Study Group. *Ann. Neurol.* 48, 927-931

23 Dessa,S.A. *et al.* (2001) Recurrence risks to sibs of MS index cases: impact of consanguineous matings. *Neurology* 56, 784-785

24 Jersild,C. *et al.* (1973) HL-A antigens and diseases. I. Multiple sclerosis. *Tissue Antigens* 3, 243-250

25 Naito,S. *et al.* (1972) Multiple sclerosis: association with HL-A3. *Tissue Antigens* 2, 1-4

26 Olerup,O. and Hillert,J. (1991) HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. *Tissue Antigens* 38, 1-15

27 Harbo,H.F. *et al.* (2004) Genes in the HLA class I region may contribute to the HLA class II-associated genetic susceptibility to multiple sclerosis. *Tissue Antigens* 63, 237-247

28 Coraddu,F. *et al.* (1998) HLA typing in the United Kingdom multiple sclerosis genome screen. *Neurogenetics.* 2, 24-33

29 Oturai,A.B. *et al.* (2004) Concordance for disease course and age of onset in Scandinavian multiple sclerosis coaffected sib pairs. *Mult. Scler.* 10, 5-8

30 Marrosu,M.G. *et al.* (2001) Dissection of the HLA association with multiple sclerosis in the founder isolated population of Sardinia. *Hum. Mol. Genet.* 10, 2907-2916

31 Coraddu,F. *et al.* (1998) HLA associations with multiple sclerosis in the Canary

Islands. *J. Neuroimmunol.* 87, 130-135

32 Saruhan-Direskeneli,G. *et al.* (1997) HLA-DR and -DQ associations with multiple sclerosis in Turkey. *Hum. Immunol.* 55, 59-65

33 Fogdell-Hahn,A. *et al.* (2000) Multiple sclerosis: a modifying influence of HLA class I genes in an HLA class II associated autoimmune disease. *Tissue Antigens* 55, 140-148

34 Haines,J.L. *et al.* (1996) A complete genomic screen for multiple sclerosis underscores a role for the major histocompatability complex. The Multiple Sclerosis Genetics Group. *Nat. Genet.* 13, 469-471

35 Sawcer,S.J. *et al.* (2004) Enhancing linkage analysis of complex disorders: an evaluation of high-density genotyping. *Hum. Mol. Genet.* 13, 1943-1949

36 Ebers,G.C. *et al.* (1996) A full genome search in multiple sclerosis. *Nat. Genet.* 13, 472-476

37 Akesson,E. *et al.* (2002) A genome-wide screen for linkage in Nordic sib-pairs with multiple sclerosis. *Genes Immun.* 3, 279-285

38 Harbo,H.F. *et al.* (2003) Two genome-wide linkage disequilibrium screens in Scandinavian multiple sclerosis patients. *J. Neuroimmunol.* 143, 101-106

39 A meta-analysis of whole genome linkage screens in multiple sclerosis. J**.** Neuroimmunol, 2003. 143(1-2):p. 39-46

40 Sawcer,S. and Compston,A. (2003) The genetic analysis of multiple sclerosis in Europeans: concepts and design. *J. Neuroimmunol.* 143, 13-16

41 Dai,K.Z. *et al.* (2001) The T cell regulator gene SH2D2A contributes to the genetic susceptibility of multiple sclerosis. *Genes Immun.* 2, 263-268

42 Smerdel,A. *et al.* (2004) Genetic association between juvenile rheumatoid arthritis and polymorphism in the SH2D2A gene. *Genes Immun.* 5, 310-312

43 Spurkland,A. *et al.* (1998) Molecular cloning of a T cell-specific adapter protein (TSAd) containing an Src homology (SH) 2 domain and putative SH3 and phosphotyrosine binding sites. *J. Biol. Chem.* 273, 4539-4546

44 Kristiansen,O.P. *et al.* (2000) CTLA-4 in autoimmune diseases--a general susceptibility gene to autoimmunity? *Genes Immun.* 1, 170-184

45 Teutsch,S.M. *et al.* (2004) Association of common T cell activation gene polymorphisms with multiple sclerosis in Australian patients. *J Neuroimmunol.* 148, 218-230

46 Fukazawa,T. *et al.* (1999) CTLA-4 gene polymorphism may modulate disease in

Japanese multiple sclerosis patients. *J. Neurol. Sci.* 171, 49-55

47 Maurer,M. *et al.* (2002) CTLA4 exon 1 dimorphism is associated with primary progressive multiple sclerosis. *J Neuroimmunol.* 131, 213-215

48 Kantarci,O.H. *et al.* (2003) CTLA4 is associated with susceptibility to multiple sclerosis. *J Neuroimmunol.* 134, 133-141

49 Masterman,T. *et al.* (2002) CTLA4 dimorphisms and the multiple sclerosis phenotype. *J Neuroimmunol.* 131, 208-212

50 Ueda,H. *et al.* (2003) Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. *Nature* 423, 506-511

51 Amundsen,S.S. *et al.* (2004) Genetic analysis of the CD28/CTLA4/ICOS (CELIAC3) region in coeliac disease. *Tissue Antigens* 64, 593-599

52 Blomhoff,A. *et al.* (2004) Polymorphisms in the cytotoxic T lymphocyte antigen-4 gene region confer susceptibility to Addison's disease. *J Clin. Endocrinol. Metab* 89, 3474-3476

53 Saiki,R.K. *et al.* (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350-1354

54 Fischer,S.G. and Lerman,L.S. (1983) DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. *Proc. Natl. Acad. Sci. U. S. A* 80, 1579-1583

55 Bjorheim,J. *et al.* (2003) Approach to analysis of single nucleotide polymorphisms by automated constant denaturant capillary electrophoresis. *Mutat. Res.* 526, 75-83

56 Ekstrom,P.O. *et al.* (2002) Population screening of single-nucleotide polymorphisms exemplified by analysis of 8000 alleles. *J Biomol. Screen.* 7, 501-506

57 Minarik,M. *et al.* (2003) Cycling gradient capillary electrophoresis: a low-cost tool for high-throughput analysis of genetic variations. *Electrophoresis* 24, 1716-1722

58 Harbo,H.F. *et al.* (1999) CTLA4 promoter and exon 1 dimorphisms in multiple sclerosis. *Tissue Antigens* 53, 106-110

59 Colhoun,H.M. *et al.* (2003) Problems of reporting genetic associations with complex outcomes. *Lancet* 361, 865-872

60 Helgason,A. *et al.* (2005) An Icelandic example of the impact of population structure on association studies. *Nat. Genet.* 37, 90-95

61 Bonetti,A. *et al.* (2004) A two-stage study on multiple sclerosis susceptibility and chromosome 2q33. *Genes Immun.* 5, 142-146

62 Dyment,D.A. *et al.* (2002) No evidence to support CTLA-4 as a susceptibility gene in MS families: the Canadian Collaborative Study. *J Neuroimmunol.* 123, 193-198

63 Ligers,A. *et al.* (2001) CTLA-4 gene expression is influenced by promoter and exon 1 polymorphisms. *Genes Immun.* 2, 145-152

64 Maurer,M. *et al.* (2002) A polymorphism in the human cytotoxic T-lymphocyte antigen 4 (CTLA4) gene (exon 1 +49) alters T-cell activation. *Immunogenetics* 54, 1-8

11 PAPERS